Impact of an added milk intervention to a micronutrient fortified school feeding program on child diet, cognition, learning, nutritional status, school attendance, and health: Statistical Analysis Plan

1. Administrative information

Version history

v.1 Date: 2024-08-27

Authors of the statistical analysis plan (SAP):

Primary author: Lilia Bliznashka[1]

Contributing authors: Aulo Gelli[1]

Affiliations:

1. International Food Policy Research Institute

Funding Source: CGIAR Initiative on Fragility, Conflict, and Migration; HSA Group

Study Registration: ISRCTN12225603

This SAP is based on the methods of analysis described in the study protocol approved by the IFPRI IRB (NDH-23-0942, approved 28/9/2023) and posted on the study registration site on May 22, 2024.

2. Introduction

2.1 Background and Rationale

Even before the onset of the current civil war, Yemen was one of the poorest countries in the world ranking 183 out of 191 countries and territories on the UN Human Development Index. The World Food Programme (WFP) estimates that currently 17 million Yemeni (half the population) are food insecure and 2.2 million preschool children are acutely malnourished (World Food Programme, 2023a). In 2022, 17% of school children in Yemen received school meals (WFP, 2022) and in 2023, the humanitarian response will cover only 8% of the needs of education sector, leaving it with the second highest unmet need (OCHA, 2023). WFP provides nutritious snacks (imported or locally procured), either fortified date bars or fortified high energy biscuits, to 1.55 million school children. There is an urgent need to understand how to improve access to nutritious school meals to support students and schools throughout the country.

2.2 Objectives and hypotheses

The aim of this study is to evaluate the effectiveness and cost-efficacy of adding a daily drink of milk to an ongoing school feeding program to improve children's diet.

3. Study methods

3.1 Trial design

In this cluster randomized controlled effectiveness trial, 42 schools in Al Mhuka district, Taiz Governorate, Yemen were randomly assigned to one of two groups: (1) Standard of care (control) group, which WFP's school feeding program involving daily distribution of high energy biscuits (n=20); and (2) School milk (intervention) group, which received the standard of care intervention with the addition of a daily drink of milk. A baseline survey was conducted prior to the start of the milk distribution in November 2023. An endline survey was conducted at the end of the school year in May 2024.

3.2 Recruitment

The primary reference group for this study is primary school aged children enrolled in schools supported by the school feeding program. A secondary reference group includes their caregivers. Lists of currently enrolled children were obtained from each school at the start of the school year. Children were randomly selected for participation in the survey, where the selection was stratified by gender.

3.3 Eligibility (inclusion / exclusion criteria)

Inclusion criteria

- Children enrolled at baseline in schools involved in the study
- Adult caregivers (≥18 years of age) of children included in the study

Exclusion criteria

- Household head, child, parent or guardian unwilling to participate in the study

The original intent to limit eligibility to children aged 6-8 y at baseline was dropped at enrolment due to the small number of children available in this age group in each school.

3.4 Randomization

A list including 42 eligible schools in the targeted districts was obtained from the Ministry of Education by in-country implementation partners. Data on school enrolment was used to classify the targeted schools into tertiles of school size. The 42 schools (clusters) were then allocated to the two study groups using a simple randomisation procedure performed in Stata, with randomisation stratified by school size tertile. The original intent of using restricted randomisation to allocate schools to interventions was abandoned due to the unavailability of school level data prior to the baseline and the need to allow implementers to plan for implementation roll-out as soon as the baseline survey was completed.

3.5 Sample size, power, and detectable difference

Power calculations based on available clusters in targeted districts and resource availability suggested 20 clusters (schools) per intervention arm and 30 households (with index children) per cluster. The primary outcomes of the trial include the 10-food group dietary diversity score in primary school children. For this outcome, assuming an inter cluster correlation coefficient (ICC) of 0.05, a sample size of 30 children per school leads to a minimum detectable effect size (MDES) of 0.26 SDs.

4. Sample characteristics

4.1 Baseline characteristics

The following variables will be presented to describe the characteristics of the study sample. They will be presented by intervention arm.

Variable	Definition
Child age	Age in years
Child sex	Male vs female
Child grade	Current grade in school
Caregiver age	Age in years of primary caregiver
Caregiver education	Level of education of primary caregiver: no formal education complete, primary complete, secondary complete, higher complete. Depending on the distribution we will also create an indicator for any formal education vs none.
Household head age	Age in years of head of household
Household head occupation	Primary occupation of the head of household
Household wealth quintile	Wealth quintile, derived from a wealth index, which will be derived hosing characteristics and asset ownership using Principal Components Analysis
Household size	Total number of household members

4.2 Analysis of attrition

Attrition rates will be presented by intervention arm using a CONSORT diagram. We will test for differences between attrited and non-attrited children at baseline using t-tests in terms of the characteristics described in 4.1 above. T-tests will be considered significant at $p<0.05$. If significant differences are found, we will perform attrition analyses using inverse probability weighting.

5. Adherence, protocol deviations, and analysis sample

5.1 Adherence

Adherence was monitored through a self-reported receipt of the school meals during a reference week. Adherence will not be defined as receiving the school meals five times during the reference week (once per day).

5.2 Protocol deviations

No protocol deviations were reported during the trial.

5.3 Definition of analysis sample

The sample will be analysed using an intention-to-treat approach.

6. Outcome variables

6.1 Data collection procedures

Data were collected using quantitative household questionnaires administered to the child and primary caregiver.

6.2 Outcome variable definitions

- Diet diversity score in children will be measured using the Minimum dietary diversity for women (MDD-W) guidelines at baseline and endline. We will calculate a count score, summing 10 foods groups, and a binary indicator for consuming ≥5 food groups in the previous 24h. Milk consumption will be calculated using self-reported consumption in the previous 24h.
- Cognition in children will be measured using the forward and backward digit span and the standard progressive matrices at baseline and endline. We will calculate raw scores summing the correct responses to each module. We will also calculate standardized scores, standardized within the sample. If age and sex distributions allow, we will calculate age and sex standardized scores. We will also use structural equation modeling to examine cognition as a latent construct.
- Learning in children will be measured using literacy and numeracy scores at baseline and endline. We will calculate raw scores by summing the correct responses to each set of questions. We will also calculate standardized scores, standardized within the sample. If age and sex distributions allow, we will calculate age and sex standardized scores.
- Nutritional status will be measured using body mass index Z-score (BMIZ) and height-for-age Z-score (HAZ) at baseline and endline. BMIZ and HAZ will be calculated using the WHO 2007 Growth Reference (2). Stunting will be defined as HAZ <-2. Thinness will be defined as BMIZ <-2. Overweight will be defined as BMIZ as >1 and obesity as BMIZ >2 (2).
- School attendance will be measured using school observations at baseline and endline. We will calculate the total number of students present on the day of the visit.
- Perceptions of the school feeding program will be measured using caregiver self-report at baseline and endline. Binary indicators will be created for different variables to show the proportion of caregivers with a perception.
- Child health will be measured using caregiver report of child morbidity symptoms at baseline and endline. We will create binary indicators for each symptom experienced. We will also calculate a summary indicator for the total number of symptoms experienced.

6.3 Primary outcome

Individual daily diet diversity and milk consumption are the primary outcomes[1]. Cognition, learning, nutritional status, attendance, perceptions, and health are secondary outcomes.

[1] There was an inconsistency with pre-specified primary research question (including both dietary diversity and nutritious food consumption) and primary outcome specification in trial registration (only including dietary diversity). This wording resolves consistency and aligns with primary research question.

7. Covariates and interaction terms

7.1 Data collection procedures

Data were collected using quantitative household questionnaires administered to the child and primary caregiver.

7.2 Covariate variable definitions

Adjusted analyses will control for the following confounders, all of which were selected a priori. Covariates will be defined the same way as described in section 4.1

Variable	Definition
Child age	Age in years
Child sex	Male vs female
Child grade	Current grade in school
Caregiver age	Age in years of primary caregiver
Caregiver education	Level of education of primary caregiver: no formal education complete, primary complete, secondary complete, higher complete. Depending on the distribution we will also create an indicator for any formal education vs none.
Household head age	Age in years of head of household
Household head occupation	Primary occupation of the head of household
Household wealth quintile	Wealth quintile, derived from a wealth index, which will be derived hosing characteristics and asset ownership using Principal Components Analysis
Household size	Total number of household members

7.3 Effect modification testing or subgroup analysis

Effect modification will be assessed by interacting treatment assignment with the following potential effect modifiers, selected a priori: child sex (male vs female), child age (continuous, categorical for each of the following groups 6-9, 10-14, 15-18 y), caregiver age (continuous, categorical for <20, 20-29, 30-39, 40-49, etc depending on the distribution, caregiver education (any formal vs none, to increase power), and household wealth (bottom two vs top three, to increase power). Effect modification will be considered significant at $p<0.10$.

8. Statistical analysis

8.1 Framework, confidence intervals and p-values

All tests will be two-sided, at 95% confidence interval and a 5% significance level unless otherwise indicated (i.e., interaction analyses). No adjustments will be made for multiple hypothesis testing (3).

8.2 Missing data

Missing data on any covariates will be imputed using mean cluster imputation if <5% of data is missing and multiple imputation if ≥5% of missingness. Missing data on outcomes will not be imputed.

8.3 Statistical analysis plan for each hypothesis

We will a single difference model specification of the following form:

$$Y_{i1} = \beta_0 + \beta_1 T_i + \beta_2 Y_{i0} + \varepsilon_i$$

where Y_{i0} is the outcome variable at baseline, $Y_{i1}$is the outcome variable at endline and T_i is a dummy variable for the treatment. This (ANCOVA) estimator has been shown to provide more efficient estimates of program impact than difference-in-difference estimators when autocorrelation of outcomes is low (4). To account for c-RCT design and the level of clustering of the outcome under analysis, we will employ linear multi-level regression models. The multi-level models will use both fixed effects with dummy variables for each intervention and random effects at the school level (unit of randomisation) to take into account clustering and to estimate the standard error in an unbiased manner. Alternative fixed effect models with standard errors clustered at the school level will also be considered. Primary analyses will be unadjusted for baseline covariates. In addition to the unadjusted primary analyses, we will report adjusted estimates, conditional on the covariates described in 7.2 above. We will use linear multi-level regression models and estimate mean differences for continuous outcomes and risk differences for binary outcomes. Alternative log-binomial and log-Poisson models will be considered for binary outcomes.

The same model will be repeated with the addition of an interaction term to assess effect modification for the variables listed in 7.3 above.

8.5 Statistical assumptions and diagnostics

All variables will be assessed for normality using a Shapiro-Wilk test. If a variable is skewed, it will be transformed via log transformation. We will assess collinearity using pairwise Pearson correlations between all covariates. If two covariates are collinear with one another, one will be selected for inclusion in the regression.

8.6 Statistical software

Statistical analyses will be conducted in Stata 18.

9. Brief description of modifications to the SAP

This is the first version of the SAP, and it has not been modified.

10. References

1. Hayes RJ, Moulton LH. Cluster Randomised Trials, Second Edition. 2nd Editio. Chapman and Hall/CRC; 2017.
2. WHO. WHO Growth Reference Data for 5–19 Years [Internet]. 2007. Available from: https://www.who.int/toolkits/growth-reference-data-for-5to19-years
3. Leroy JL, Frongillo EA, Kase BE, Alonso S, Chen M, Dohoo I, et al. Strengthening causal inference from randomised controlled trials of complex interventions. BMJ Glob Heal. 2022 Jun 10;7(6):e008597.
4. Bruhn M, McKenzie D. In Pursuit of Balance: Randomization in Practice in Development Field Experiments. Am Econ J Appl Econ. 2009 Sep 1;1(4):200–32.